Unemployment Insurance, Job Search, and Informal Employment

David Margolis, Lucas Navarro, David A. Robalino

Unemployment Insurance, Job Search and Informal Employment

David N. Margolis*, Lucas Navarro† and David A. Robalino‡

June 14, 2012

Abstract

This paper analyses the potential impacts of introducing unemployment insurance (UI) in middle income countries using the case of Malaysia, which today does not have such a system. The analysis is based on a job search model with unemployment and three employment sectors: formal and informal wage employment, and self employment. The parameters of the model are estimated to replicate the structure of the labor market in Malaysia in 2009 and the distribution of earnings for informal, formal and self employed workers. The results suggest that unemployment insurance would have only a modest negative effect on unemployment if benefits are not overly generous. The main effect would be a reallocation of labor from wage into self employment while increasing average wages in the formal and informal sectors.

Keywords: Unemployment insurance, Informal sector, Self employment, Job search

JEL Codes: J64, J65, O17, J23, J31, J21, J62

*The World Bank

†ILADES-Universidad Alberto Hurtado

‡The World Bank

This paper was prepared as part of a set of Analytical and Advisory Activities provided to the Government of Malaysia. Access to the data used here was provided under conditions set by Malaysia's Department of Statistics and the Economic Planning Unit; interested individuals are invited to contact the DoS and EPU to arrange access if desired.

1 Introduction

In developing countries, income protection in the case of unemployment is mainly provided through severance pay, which is regulated by the labor code. These systems cover only a minority of the labor force, are difficult to enforce, and carry considerable risk of default, particularly during times of crisis, given that employers seldom provision for the benefits (Holzmann et al., 2011). Meanwhile, only a few countries have introduced contributory unemployment benefit programs. Part of the reason is the concern that traditional unemployment insurance systems (or providing basic unemployment assistance) can lead to abuse, promote unemployment, and generate large fiscal outlays, particularly in the presence of informal employment and weak institutional capacity to monitor job-search and employment status.

The empirical evidence, mainly coming from high income countries, is mixed.[1] In general, studies find that there is a positive correlation between the level and duration of benefits and the length of the unemployment spell, which can lead to higher unemployment rates. At the same time, evidence from Denmark, France, Germany, Italy, Ireland and Spain shows that longer unemployment spells can also be associated with "better matches"; workers are able to find better and more stable jobs.(Tatsiramos, 2009)

Less is known about the effects of unemployment insurance in developing countries. The few studies available would suggest that effects on the duration of unemployment spells and employment levels are modest. Two studies for Brazil find no quantitatively meaningful effect of the UB system on the duration of unemployment spells. The first exploited changes in eligibility conditions and showed that, if anything, unemployment benefits allowed faster transitions into self employment (Cunningham, 2000). A second compared the exit rates from unemployment among formal sector workers (eligible for unemployment benefits) and informal sector workers (not eligible for unemployment benefits). It showed that the former had higher exit rates even after controlling for unobserved characteristics that are correlated with work in the formal sector (Margolis, 2008). At the other extreme, van Ours and Vodopivec (2008) show that in the case of Slovenia, the shortening of the potential duration of UI benefits substantially reduced the length of the unemployment spell. A more recent paper based on a structural model for Brazil finds the unemployment insurance system mainly reduces transitions into informal jobs (Robalino et al., 2011).

This paper analyses the potential impacts of introducing an unemployment insurance (UI) system in Malaysia, which currently has no such system. The analysis is based on a structural job search model with unemployment and three employment sectors: formal and informal wage employment, and self employment. The parameters of the model are estimated to replicate the structure

[1] For reviews see Holmlund (1998); Vodopivec et al. (2005); and Olinto et al. (2007).

of the labor market in Malaysia in 2009 and the distribution of earnings for informal, formal and self employed workers. The model is used to simulate the effects of alternative unemployment benefit system designs that depend on the replacement rate, the vesting period for benefits, the duration of benefits, and the contribution rate. We look at changes in the shares of individuals across employment states, the unemployment rate, and average earnings by sector.

The results suggest that an unemployment insurance system in Malaysia would have only a modest negative effect on unemployment if benefits are not overly generous. The system would induce a reallocation of labor from wage into self employment while increasing average wages in the formal and informal sectors. The effects on the average earnings of the self employed would depend on the generosity of the system. With a 50% replacement rate, most workers entering self employment would be low skilled workers, driving average earnings down. High skilled workers would change behavior significantly with more generous systems. As they remain unemployed for longer, wage employment offers fall, and a larger share enters self employment, driving up average earnings among the self employed. Although the analysis focuses on Malaysia, the results are likely to be relevant for other middle income countries, particularly in Latin America, which share similar demographic, levels of education and labor market structures.

The remainder of this paper is structured as follows. Section 2 provides a brief description of the Malaysian labor market, while 3 lays out a 4-state (unemployment, self employment, formal wage and salary employment, informal wage and salary employment) job search model . Section 4 describes the data used for the estimations, section 5 covers the reduced form and simulated pseudo method of moments estimation results, while section 6 presents the microsimulation results. Section 7 concludes.

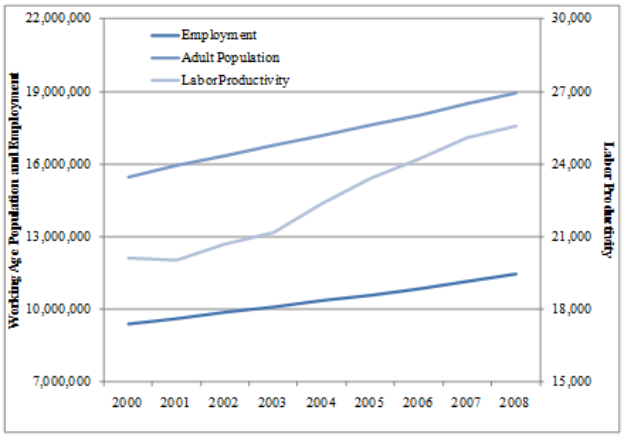
2 Overview of Current Malaysian Labor Markets

Over the last decade both employment and labor productivity have been growing in Malaysia. (see figure 1). Between 2000 and 2008 the adult population grew by an average of 2.6 percent annually over this period, while employment grew by an average of 2.8 percent per year and labor productivity grew on average by 3.7 percent.

Similar to middle income countries in Latin America, Malaysia has a young and relatively well educated working age population. Over 60 percent of the working age population is under 40 years old, and over 35 percent was under 25 years old.[2] About 57 percent of the workforce has some

[2]In the states of Sabah and Labuan, there are over 5 percent more people under the age of 25 and roughly 4 percent fewer people over the age of 55 than in the other regions. The state of Sarawak differentiates itself from Peninsular Malaysia in that its demographic bulge happened earlier, so that there are more people in their thirties and fewer people in their twenties and teens in Sarawak than in Peninsular Malaysia and Sabah/Labuan.

Figure 1: Working Age Population, Employment and Labor Productivity Trends, 2000-2008



Source: World Development Indicators.
Notes: Labor productivity is measured as GDP per worker in 1990 USD converted at PPP rates.

form of secondary education certification and 14.4 percent of the population has a post-secondary certification, with over 5 percent having a tertiary education degree. However, there are clear regional disparities; for example, the working age population in East Malaysia, especially the states of Sabah and Labuan, is much less educated than that of Peninsular Malaysia.[3]

The differences in compensation, hours worked and formality levels[4] among population groups (sex, region, age, level of education, and sector of activity) are also important (Table 2). Men are more likely to be in formal jobs (55 percent) than women (48) and also earn more – men's base monthly compensation is on average 17 percent higher than women's. Not surprisingly, education is an important determinant of labor market outcomes. Only 25 percent of those with primary education or less are in formal employment against 44 percent for those with secondary education and 68 percent for those with higher education. One explanation consistent with the model developed below is that formal contracts are not for low productivity workers, who end-up working in low productivity activities often in small firms in the informal sector or as self employed. Thus

[3]Forty percent of the working age population of Sarawak has at most a primary education, while the number rises to 47 percent for Sabah/Labuan. Conversely, nearly 60 percent of the working age population in Peninsular Malaysia has a secondary degree (compared with 52 percent in Sarawak and 43.3 percent in Sabah/Labuan) and 15.6 percent have a post-secondary certification (compared with 8.1 percent in Sarawak and nearly 10 percent in Sabah/Labuan).

[4]For the purposes of the data work in this report, employment in the "formal sector" is defined as holding a job for which the employer makes Employer's Provident Fund (EPF) or Social Security (SOCSO) contributions. By this definition, self employed individuals who are not required to contribute are considered to be"informal", and this definition will also classify some teachers or health workers employed by the public sector as informal, since their employer is not required to contribute to EPF and SOCSO on their behalf.

Table 1: Working Age Population

	Peninsular	Sarawak	Sabah and Labuan	Total
General Demographics				
Male	50.3	49.4	50.0	50.2
Malaysian Citizen	95.6	96.5	75.6	93.7
Never Married	46.8	42.7	47.3	46.5
Married	48.9	53.4	48.9	49.3
Widowed	3.3	2.7	2.6	3.2
Divorced / Separated	1.0	1.1	1.2	1.0
Highest certifcate obtained				
Not Applicable	3.6	10.2	13.1	5.1
No Certificate	10	12.8	16	10.8
UPSR/UPSRA or equivalent	11.5	17.1	17.9	12.6
PMR/SRP/LCE	23.9	24.3	20.8	23.7
SPM/MCE	35.3	27.6	22.5	33.4
STPM/HSC or equivalent	2.9	2.4	2.9	2.9
Certificate	1.5	0.8	0.9	1.4
Diploma	5.4	2.3	2.6	4.8
Degree	5.8	2.6	3.4	5.3
Status in employment				
Not Employed	49.6	46.3	46.7	49
Employer	2	1.6	1.8	1.9
Government Employee	6.8	6.2	6	6.7
Private Employee	29.7	27	30.4	29.5
Own Account Worker	9.6	12	11.5	10
Unpaid Family Worker	2.3	6.8	3.7	2.8
Age				
15-24	37.7	37.2	43.4	38.2
25-39	26.2	25.9	25.8	26.1
40-44	23.3	23.7	21.9	23.2
55-64	12.8	13.1	8.9	12.5

Source: Labor Force Survey 2009.
Note: The Labor Force Survey does not sample collective housing. All figures are percentages using LFS sample weights.

there are large difference in earnings: the average worker with higher education earns five times more than the average worker with primary education or less.

3 The Model Framework

Search and matching models have been extensively used for the analysis of the quantitative effect of labor market policy. In this paper we adapt the Diamond-Mortensen-Pissarides style of dynamic job search model (Pissarides, 2000) to fit the situation of labor markets with large informal sectors. In this type of model, unemployed individuals engage in job search activity and the decision to accept a job depends on the value of the job, which is itself a function of the job's stability and the opportunity cost of not being able to search for a better job while employed. This sort of model has been extensively analyzed and some recent extensions of this framework have introduced informal labor markets and self employment(Albrecht et al., 2009).

This paper continues in this literature by integrating both an informal sector and self employment, although it does not solve analytically for the equilibrium labor market tightness and the shares of offers from the formal and informal sector wage employers. Appendix A describes the basic structure of the model.[5] The model defines a set of value functions for different labor market states and a wage determination mechanism. The model is solved numerically in section 5.

3.1 Unemployment Value Function

Let $U(y, \tilde{w})$ be the value of unemployment for a worker of type y, which can be written as

$$\begin{aligned} rU(y,\tilde{w}) &= b(\tilde{w}) + \alpha \max\left[N_0(y) - U(y,\tilde{w}), 0\right] + \\ &\quad m(\theta)\psi E \max\left[N_I(x,y) - U(y), 0\right] + \\ &\quad m(\theta)(1-\psi) E \max\left[N_F(x,y) - U(y,\tilde{w}), 0\right]. \end{aligned} \tag{1}$$

This worker receives a flow utility of $b(\tilde{w})$, where $\tilde{w}$ is the wage on the previous job, while unemployed.[6,7] At a rate α, the worker meets an opportunity for self employment and, if it is taken,

[5] See Margolis et al. (2011) for the full structural model with expressions for equilibrium employment rates in each sector and wage distributions.

[6] For simplicity of exposition, it is assumed that all ex-workers, including self employed and informal sector workers, can draw unemployment benefits, as a function of their earnings prior to unemployment. This assumption is relaxed in the numerical simulations, where the replacement rate for unemployment spells following self employment or informal wage employment is set to zero.

[7] It is assumed that individuals can only receive unemployment benefits when unemployed, i.e. not when self employed or in informal employment. This implies significant enforcement capacity on behalf of he government, which may not be appropriate. We discuss the implications that relaxing this assumption might have in section 6.

Table 2: Compensation, Hours, Employment and Formality

	Base Monthly Compensation	Total Monthly Compensation	Total Weekly Hours Worked	Total Hourly Compensation	Share of Employment	Share Employer/ Own-Account	Share Formal: All Jobs	Share Formal: Wage & Salary
Women	1,246	1,331	44.0	7.41	36.1%	15.1%	54.9%	68.0%
Men	1,456	1,580	46.0	8.51	63.9%	28.1%	47.7%	61.5%
Peninsular	1,484	1,596	45.1	8.69	80.8%	23.0%	51.8%	64.2%
Sarawak	1,012	1,094	44.9	6.13	9.0%	25.5%	42.3%	62.0%
Sabah and Labuan	821	938	47.2	4.94	10.1%	24.8%	42.6%	62.0%
Age 15-24	790	854	46.1	4.55	19.1%	8.1%	61.9%	67.8%
Age 25-39	1,514	1,635	45.2	8.86	39.4%	16.0%	59.9%	69.7%
Age 40-54	1,705	1,844	44.8	10.14	31.5%	31.0%	41.1%	57.3%
Age 55-64	1,442	1,548	44.5	8.64	10.0%	58.1%	16.7%	35.8%
Completed Primary or Less	802	862	46.1	4.55	29.1%	36.1%	25.1%	43.3%
Completed Lower Secondary	987	1,059	46.0	5.60	14.2%	26.9%	44.1%	58.1%
Completed Upper Secondary	1,163	1,274	45.5	6.87	35.4%	17.9%	58.3%	68.8%
Completed Pre-University	1,952	2,125	43.6	11.89	12.5%	13.8%	65.2%	72.2%
Completed Higher Education	4,111	4,358	42.7	24.29	8.7%	11.5%	68.1%	73.4%
Resource-Based Industry	952	1,027	44.7	5.62	6.9%		15.3%	44.1%
Manufacturing	1,355	1,454	44.6	7.93	31.1%		77.8%	84.0%
Services	1,668	1,952	43.7	11.36	62.0%		48.3%	58.6%

Source: Labor Force Survey 2009 (compensation, hours and employment share); Household Income Survey 2007 (Share of formal jobs).
Note: Base Monthly Compensation, Total Monthly Compensation, and Total Hourly compensation are expressed in 2010 RM by using the CPI to convert the nominal 2009 figures. Figures refer only to wages and salaries received and do not include any residual income earned by entrepreneurs. Calculations include private employees and own-account workers for employment share, hours, and compensation.

realizes a capital gain of $N_0(y) - U(y, \tilde{w})$. The matching function $m(\theta) = a\theta^{\beta}$ determines the probability of an individual meeting a vacancy for a wage and salary job, with θ being the ratio of vacancies to unemployment (i.e. a measure of labor market tightness). Conditional on meeting a vacancy, a worker meets a formal sector vacancy with probability $1 - \psi$ and an informal sector vacancy with probability ψ. If the job is taken, the worker realizes a capital gain of $N_j(x, y) - U(y, \tilde{w})$. The employment value functions consider that the initial flow value of the match is given by the sector-specific productivity x, which varies across individuals and sectors (formal and informal).

3.2 Self Employment Value Function

Let $N_0(y)$ be the value of self employment for a worker of type y, which can be written as

$$rN_0(y) = y + \lambda_0 (U(y, y) - N_0(y)). \tag{2}$$

This expression shows that the self employed receives a flow value equivalent to her type y, but at rate λ_0 the opportunity ends in which case there is a (negative) capital gain of $U(y, y) - N_0(y)$.[8]

3.3 Wage and Salary Sector Value Functions

Let $N_F(x, y)$ and $N_I(x, y)$ be the values of employment in the formal sector and informal salaried sector, respectively, for a worker of type y with a draw x from the sector-specific productivity distribution. The flow values of these jobs can be written as

$$rN_F(x, y) = w_F(x, y) + \lambda_F (U(y, w_F(x, y)) - N_F(x, y)) \tag{3}$$

$$rN_I(x, y) = w_I(x, y) + \lambda_I (U(y, w_I(x, y)) - N_I(x, y)), \tag{4}$$

respectively. A worker of type y who has a formal sector job receives a wage $w_F(x, y)$ determined by Nash bargaining. Idiosyncratic shocks that destroy the match and send the individual back to unemployment arrive at rate λ_j. A similar idea applies to the value of employment in the informal wage sector.

[8] As savings are not a component of the model, it is assumed that taking a self employment opportunity requires no startup capital.

3.4 Employer-Side Value Functions

Let V_j and $J_j(x,y)$ be the value of the expected profit of posting a vacancy and the value of a filled job in sector $i \in \{F, I\}$. The filled job values can be written as

$$\begin{aligned} rJ_F(x,y) &= x - w_F(x,y)(1+\tau) + \lambda_F(V_F - J_F(x,y) - s) \\ rJ_I(x,y) &= \epsilon x - w_I(x,y) + \lambda_I(V_I - J_I(x,y)). \end{aligned}$$

Regulations affecting the formal sector filled job value are payroll taxes τ and severance payments s. Note that regulations do not affect $J_I(x,y)$ but the flow value of productivity therein is ϵx, where $\epsilon < 1$.

The values of formal and informal sector vacancies, V_F and V_I, are defined respectively by

$$\begin{aligned} rV_I &= -c + \frac{m(\theta)}{\theta} E \max[J_I(x,y) - V_I, 0] \\ rV_F &= -c + \frac{m(\theta)}{\theta} E \max[J_F(x,y) - V_F, 0]. \end{aligned} \tag{5}$$

The expectation term in (5) reflects the assumption that the firm does not know in advance which type of worker it will meet.

3.5 Wages

As standard in this literature, a surplus is realized when a match is formed. This surplus is given by the net gain from matching for both the firm and the worker, that is, $N_s(x,y) - U(y) + J_s(x,y) - V_s$. Wages are determined by rent sharing over the surplus of the match. A wage is a solution to a generalized Nash Bargaining problem with threat points equal to the worker's and the firm's respective continuation values. Given an exogenous share parameter β and the free entry condition ($V_j = 0$), the formal contract wage $w_F(x,y)$ for a worker of type y producing at x solves

$$\max_{w_F(x,y)} [N_F(x,y) - U(y)]^{\beta} [J_F(x,y) + s]^{1-\beta}$$

which implies,

$$w_F(x,y) = \beta \frac{(x + rs)}{1+\tau} + (1-\beta) r U(y). \tag{6}$$

Similarly, the informal wage $w_I(x,y)$ solves

$$\max_{w_I(x,y)} [N_I(x,y) - U(y)]^{\beta} J_I(x,y)^{1-\beta},$$

and the wage function can be written as

$$w_I(x,y) = \beta \epsilon x + (1-\beta) r U(y). \tag{7}$$

The wage in sector i is the weighted average of the worker's productivity and the worker's outside options. Of course, regulations affect the formal wage bargaining process.

4 The Data

The model is estimated on two separate data sources from Malaysia, the Household Income Survey,[9] or HIS, and the Labor Force Survey with the Wages and Salaries Supplement,[10] or LFS. These data sources are complementary in that neither source is perfectly suited for estimation of the model presented in section 3, but the limits of each source can be compensated by the advantages of the other. Both data sources are thus used to estimate different components of the model, and as both surveys are designed to be representative of the same population (and sample weights are always used), no further correction for sampling is necessary[11].

4.1 The Household Income Survey (2009)

The HIS is a household survey that is representative of the population of Malaysia that does not live in collective housing.[12] It is comprised of 4 components: basic identification, sociodemographic characteristics of the household members, basic information on activity status and sources of income.

The HIS is particularly useful for the purposes of this paper because it allows for separate measurement of income from paid employment and self employment income. In terms of income from paid employment, the survey distinguishes between the various components and separately measures employer contributions to social security, pensions, etc. on the worker's behalf.[13] It does not, however, directly measure the amount received in severance payments.[14] In terms of self employment income, it allows for the separation of total self employment earnings into own consumption use and income, with the latter being broken into agricultural and non-agricultural self employment earnings.

There are two main difficulties with using exclusively the HIS for this paper. First, the information on labor market status does not specifically designate unemployment as a labor market state. The available categories are:

[9]The HIS is collected by the Department of Statistics for the Economic Planning Unit.

[10]The LFS is collected by the Department of Statistics.

[11]It should be noted that some variables needed to be built form both sources. See section 5.1 for details.

[12]The exclusion of collective housing is likely to be an issue for some sectors, particularly agriculture where many workers live in collective housing in the middle of plantations.

[13]The actual question asked is "How much did ..... earn during the last twelve (12) months from paid employment?' for which one earnings category is "Employer's contributions to EPF, SOCSO, etc.".'

[14]It should be noted that Malaysia does not have an explicit unemployment insurance system, however it does provide for mandatory severance pay in the case of dismissal or redundancy-based separation from (formal) jobs (Holzmann and Vodopivec, eds, 2011).

1. Employer
2. Government employee
3. Private employee
4. Own account worker
5. Unpaid family worker
6. Housewife / Looking after home
7. Student
8. Child not at school
9. Others (specify)

It is therefore impossible to calculate the share of the work force in each state as defined by out model, unless one assumes that all codes other than 1-4 are unemployed.

Second, the survey collects data referring to the previous year's values, so the questions on activity status are phrased as

> What did ..... (usual member) ..... do MOST OF THE TIME during the last twelve (12) months?

and the income questions refer to all income received during the previous year. There is no calendar, so it is impossible to know whether a low amount of earnings is due an individual working part of the year or being poorly paid for a full year of work. Moreover, and more importantly for the estimation of the model in section 3, it is impossible to measure the length of employment and unemployment spells.

4.2 The Labor Force Survey with the Wages and Salaries Supplement (2009)

The LFS is a standard labor force survey, focusing on the working age population that does not live in collective housing. The LFS has 3 components: identification, household member characteristics and labor force particulars. The wages and salaries supplement provides additional information on income for government and private employees only.

The main advantage of the LFS over the HIS is that its "labor force particulars" component is more thorough than that of the HIS. Questions are asked relative to the reference week, ILO-standard unemployment definitions are used, and there is enough information to calculate the distribution of unemployment durations within the stock of unemployed.[15]

Unfortunately, the income component of the LFS is not as rich as the HIS. In particular, there is no information on earnings from any source other than wages, salaries and overtime payments. This implies that the data cannot be used for estimating y directly, as self employment earnings cannot

[15] Unemployment durations are measured as: less than 3 months, 3 months - less than 6 months, 6 months - less than 1 year, 1 - 3 years and more than 3 years.

be seen. Moreover, the wages and salaries supplement measures gross earnings but provides no means of quantifying the value of employer or employee contributions to social insurance, implying that one cannot distinguish formal from informal wage and salary work from this data source.

5 Estimation

The structural parameters of the model were estimated by a simulated pseudo-method of moments procedure as follows. A reduced form model was estimated to recover values of y for each individual in the population, as well as the distributions $G_F(x|y)$ and $G_I(x|y)$. With this information, careers were simulated as follows. Starting from unemployment, the individual receives self employment offers, formal job offers and informal job offers, and accepts or refuses them according the behavior dictated by the value functions in section 3. When self employed, formally employed or informally employed, job destruction shocks arrive with probabilities λ_0, λ_F and λ_I respectively. Each individual's transitions and wages were repeatedly simulated until the simulations converged to a stable share of the population in each labor market state and stable wage distributions within the formal and informal employment states. The structural parameters of the model were adjusted until the steady state shares in each labor market state and the steady state wage distributions matched those found in the data, at which point the model was considered to have converged.

It is important to stress that the full structural model was not estimated, meaning that the parameters and the distribution of y variables recovered is robust to certain functional form assumptions.[16] The simulations do exploit the functional forms of expressions 1, 2, 3 and 4, and the productivity draws are assumed to come from logistic distributions. However, the steady state version of the model is not solved analytically nor is it structurally estimated.[17] It is also worth noting that the income variables used in the estimation are those drawn from a data source that reflects annual incomes rather than monthly, weekly or hourly earnings. The absence of data on earnings in self employment prevented us from using the Malaysian LFS for all individuals, and mixing data sources for measurement of a single variable is not good practice.

The remainder of this section describes how each of the steps was performed.

5.1 Direct recovery of parameters and population moments

There are several key parameters of the model that can be directly recovered from the data.

[16] This is not meant to imply that the estimation is fully non-parametric; in particular, a Heckman-type selection correction is employed in the estimation of y, implying joint normality between the disturbance term of the expression that determines Y and that which determines the probability of self employment; see section 5.2 for details.

[17] See Margolis et al. (2011) for the full structural estimation.

- τ, *the payroll tax rate for workers in the formal sector.* As mentioned in section 4.1, the Household Income Survey (HIS) data allows for separate measurement of gross compensation and and employer contributions to social insurance. It is therefore straightforward to calculate

$$\tau = \frac{\text{Total Compensation Cost}}{\text{Gross Compensation}} - 1 = 0.123.$$

- *Share of population in each employment state (Unemployment, Formal Employment, Informal Employment and Self Employment).* As mentioned in section 4, the HIS does not allow for the identification of unemployment while the Labor Force Survey (LFS) does not allow one to distinguish formal from informal wage and salaried employment. However, the estimation strategy only requires the shares of the active population in each employment state, not the identification of the employment state for any particular individual. Accordingly, the shares in unemployment (0.0202), self employment (0.2140) and wage and salary employment (0.76857) were calculated directly from the LFS. The share of wage and salary employment that is in the formal sector (0.6451) was calculated directly from the HIS. This latter share was multiplied by the share in wage and salary employment to obtain the population share in formal employment (0.4940) and the share in informal employment (0.2717).

- *Mean and variance of log earnings in the formal, informal and self employment sectors.* These values were calculated directly from the HIS data, with the following results.

Table 3: Moments of Earnings Distributions, by Sector

Sector	Mean	Standard Deviation
Self Employed	9.436	1.069
Formal	10.001	0.716
Informal	9.640	0.901

5.2 Recovery of y for All Sampled Individuals

According to the model, individuals in self employment earn the value of their type, y_i. This implies that one can use observed income of those in self employment to characterize the determinants of an individual's type, and use this estimation to recover an estimate of the (unobserved) type for each person in the rest of the population, namely the unemployed and those in wage and salary employment.

Worker type is assumed to be a function of observed characteristics[18] Z_i, i.e.

$$y_i = Z_i\gamma \tag{8}$$

However, we further assume that there can be idiosyncratic shocks that affect the income an individual draws from her self employment at a point in time, and these shocks can be written as $\nu \sim N(0, \sigma_\nu)$. This implies that the income equation from self employment takes on a standard Mincerian form, namely

$$w_i^{SE} = y_i + \nu_i = Z_i\gamma + \nu_i \tag{9}$$

As suggested by section 3, being observed in employment is the result of an optimization decision conditional on benefits received in unemployment, the income drawn from self employment, the riskiness of self employment, the likelihood of getting a formal or informal sector job offer and the wages and job stability associated with those offers. This implies that the set of individuals actually observed in self employment is a selected sample, so estimation of equation 9 requires a correction for selection bias. The selection into self employment is thus modeled as a function of observables[19] A, and we adopt the standard Heckman (1979) formulation for estimation of the vector γ. The results of this estimation are found in table 9 in the appendix. These results are used to estimate the value of of y for all individuals in the sample, including those not observed in self employment.

5.3 Estimation of the distributions of sector specific productivity draws $G_I(x\,|y)$ and $G_F(x\,|y)$

The distributions $G_j(x\,|y)$, with $j \in \{\text{Formal}, \text{Informal}\}$, are assumed to be logistic with mean equal to $y + \tilde{k}_j$, where $\tilde{k}_j$ is a form of unobserved individual-specific heterogeneity that shifts the mean of the x distribution. It is assumed that $\tilde{k}_j$ is a Bernoulli-distributed random variable. Recovering these distributions (one each for formal and informal wage employment) requires four step procedure:

1. Using the same determinants of wages as were used in the recovery of y from the self employed, estimate the part of wages that is not due to y;

2. Recover the part of this residual that is explainable by "unobserved" characteristics, and use this to characterize the distribution of the heterogeneity component $\tilde{k}_j$;

[18] The explanatory variables included in Z that determine human capital are the highest degree obtained (7 levels), sex, age, age squared and indicators for the state of residence (16 levels).

[19] The variables included in the A vector are limited by the relatively limited set of variables available in the LFS and HIS. Accordingly, the A vector includes the same variables as in Z, plus marital status (5 levels) and relation to household head (10 levels).

3. Recover the value of the draw that is added to the mean $y + \tilde{k}_j$; and

4. Calculate the distribution of these draws.

One first needs to estimate the relation between wages and y.[20] Since wages are not observed for each individual in every sector, recovery of the distributions of x begins with estimation of sector specific Mincer equations.[21] This estimation represents the first step in recovery of the unobserved heterogeneity and x distributions.

The first-stage residuals of these Mincer regressions are decomposed into a component that is explainable by the full set of variables, including the instruments that are considered "unobservables" for the purposes of the wage equation, and an orthogonal second-stage residual component. The expected first-stage residuals are divided into two groups at the median. The average predicted residual for each subgroup is calculated, the mean for the smaller half of the observations is subtracted from the expected wage resulting from the first stage regression (to normalize the support of the k_j distribution to 0 and a positive value) and their difference becomes the estimator of k_j,[22] each of which occurs (by construction) with probability 0.5.

The result of these calculations is a distribution (centered around y or $y + k_j$) of means for the draws from the x distribution. Subtracting this mean from the observed wage gives an estimator of the value of the draw from the x distribution.[23] The variance of the estimators for each distribution

[20] In the model outlined in section 3, the wage is the result of Nash bargaining over the surplus of the employment relation between the employer and the individual. The surplus is a proportional to x, and the result of the bargaining gives a share of the surplus to the individual in the form a wage that must exceed the individual's outside option. These expressions imply that the relation between the wage and x is linear; see Margolis et al. (2011) for details. Since $w = a + bx$, where a and b are unknown parameters, $E(x) = \frac{E(w)}{b} - \frac{a}{b}$ and $\text{var}(x) = \frac{\text{var}(w)}{b^2}$. Since this paper does not need to provide a structural interpretation to the intercept or slope parameters of the linear relation in order to apply the simulation estimator, one can simply recover the mean and variance of the distribution of w, which will themselves be linear functions of the underlying (and unknown) mean and variance of x, and use these estimated parameters in the simulations.

[21] Note that although selection into employment in each sector is an issue, as it was for self employment, there exist no valid exclusion restrictions for estimating a selection corrected mincer equation. This is because the wage is the result of bargaining between the worker and employer, whereas in self employment it was only a function of the individual's type and a shock. The bargaining dimension implies that wage outcome will depend on the value of the surplus, which itself is a function of all of the same factors that determine selection into one sector or the other. This dependence is modeled as affecting the distribution of $\tilde{k}_j$, namely $P\left(\tilde{k}_j = 0\right) = 1 - \Phi(A_i \kappa_j)$ and $P\left(\tilde{k}_j = k_j\right) = \Phi(A_i \kappa_j)$. The estimation of the sector specific Mincer equations can thus be written as $w_i^j = Z_i \gamma^j + \nu_i^j$. The results of these models are found in table 10 in the appendix.

[22] The estimates are $k_F = 3.167 * 10^{-3}$ and $k_I = 1.407 * 10^{-4}$.

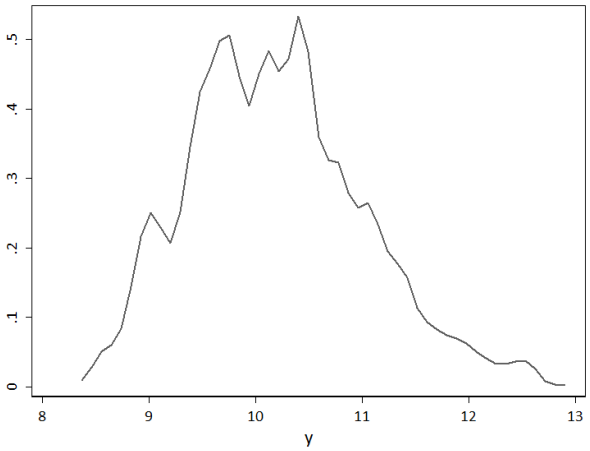
[23] More precisely, as noted in equations 6 and 7, the wage is a linear transformation of the (rescaled) reservation utility and a draw from the x distribution. As the reservation utility depends on the expected wage, which is a function of $y + k_j$, the draw of x can be approximated by the difference between the observed wage and $y + k_j$. This

is calculated and converted into a scale parameter for each sector's distribution of x draws. For the formal sector, this gives 20.726 as the variance of the logistically distributed shock, or a scale parameter for a logistic distribution of $\sqrt{\frac{3\text{var}(x)}{\pi^2}} = 2.510$ and for the informal sector this gives 20.227 as the variance of logistically distributed shock and a scale factor of $\sqrt{\frac{3\text{var}(x)}{\pi^2}} = 2.480$.

5.4 Recovery of the remaining parameters by simulated pseudo method of moments

The simulation focuses on a sample of 500 individuals drawn from the distributions of y_i and x_i[24] as estimated in the previous section, where i indexes the individual and s the sector (formal, informal, or self employment). Figure 2 describes the distribution of y in the population. At the beginning of the simulation all workers are unemployed and the number of vacancies is equal to the number of unemployed. As noted in section 3.1, the matching function $m(\theta) = a\theta^{\beta}$ determines the probability of an individual meeting a vacancy, with θ being the ratio of vacancies to unemployment (i.e. a measure of labor market tightness). Contracts are only agreed upon when both profits are positive and the flow value of a given job offer is higher than the value of unemployment.

Figure 2: Distribution of y in the Population



The micro simulation model follows over time (monthly) each of the individuals in the sample, who are characterized by y_i and x_i. Through aggregation, the model generates steady state distributions for variables of interest: the current unemployment rate in Malaysia; the shares of

approximation is best when $\beta = 1$.

[24]For consistency in the simulations, the productivity draws x are assumed to be the same across firms within the same sector for a given individual.

informal, formal, and self employment; as well as the mean and variance of the log of earnings for each type of job. To conduct policy simulations, we have to estimate 7 parameters in order to match these distributions. The parameters are: job destruction rates (3); the share of informal vacancies; the arrival rate of job opportunities for self employment; the bargaining power coefficient; and the scale factor in the matching function. The identification of the model parameters is based on the minimization of the following function in the absence of an unemployment benefit system:[25]

$$\sum_{j=1}^{3} \left(s_j - s_j^*\right)^2 + \left(w_j - w_j^*\right)^2 + \left(v_j - v_j^*\right)^2 \tag{10}$$

where s_j is the share of the labor force working in sector j, w_i is the average of log wages across individuals working in the sector, v_j is the variance, j=1: informal, 2:formal, 3:self employed, and a * indicates the targeted values of the variable, estimated from the data.

The values of the recovered parameters,[26] conditional on the sample of individuals drawn, are presented in Table 4.[27] The model is able to closely reproduce the employment shares but predicts higher wages for informal sector workers than found in the data, as seen in Table 5. We find job destruction rates ranging between 1 and 5% per month, the lowest in formal sector and the highest in the informal sector. The arrival rate of opportunities in self employment is close to 8% per month whereas the share of informal job offers is close 40%. Finally, for the sample of individuals drawn, workers appear to have a relatively high bargaining power (> 0.7) but below what would be observed in a perfectly competitive market.

6 Microsimulations of Policy Experiments

Based on the estimated parameters several policy experiments were conducted to assess the labor market impact of introducing an unemployment insurance system.

We considered 25 types of unemployment insurance systems based on variations in the values of the following policy parameters: the replacement rate; the contribution rate to the system (assumed to be paid by the employer); the vesting period (i.e., the number of contributions necessary to qualify for benefits); the duration of unemployment benefits. Although the last two are inconsistent with a literal interpretation of the model in section 3, they are nevertheless closer to situations found in the real world. The implicit assumption here is that workers are myopic and only take into account

[25] The estimation uses the ANT algorithm (Miller, 1998).

[26] The severance pay parameter, s, is set to zero for the purposes of the simulation.

[27] Clearly, different samples would lead to different values for the model parameters. In principle, one could bootstrap samples to estimate standard errors for the parameters. In the micro-simulation estimation, however, we keep the sample of individuals constant.

Table 4: Value of Model Parameters Estimated for the Policy Simulations

Parameter	Value	Description
β	0.7646	Bargaining power
a	0.5631	Scale factor in matching function
α	0.0817	Arrival of self employment opportunities
ψ	0.4	Share of informal job offers
λ_i	0.0477	Job destruction rate in the informal sector
λ_F	0.01	Job destruction rate in the formal sector
λ_I	0.01	Job destruction rate for the self employed

Table 5: Performance of Estimated Model

Population Moment	Estimated Values	Observed Value
Share informal work	0.263	0.281
Share formal work	0.538	0.511
Share self employment	0.174	0.186
Mean log-wage informal sector	10.0	9.640048
Std log wage informal	1.2	0.9005742
Mean log-wage formal sector	10.2	10.00101
Std log wage formal	0.7	0.71
Mean log-earnings self employment	9.7	9.436391
Std log-earnings self employment	1.5	1.06

the availability of unemployment benefits in the next period when deciding whether to take a given job offer. The simulations take as a reference situation a world without unemployment benefits, i.e. where the payroll tax rate, replacement rate, vesting period and benefit duration are all zero.

The ranges considered for each of the parameters are summarized in Table 6. The results of the various simulations are presented in Appendix Table 11. We also analyze the impact of the unemployment benefit system in the case of a lower value for the bargaining coefficient, chosen to be 0.3.[28] The results of these simulations are presented in Appendix Table 12. Table 7 summarizes the key results from these tables for employment shares, while table 8 summarizes the key results for average earnings.

Table 6: Values of Policy Parameters for Unemployment Benefit Systems

Policy Parameter	Set of Values
Contribution Rate	{0,0.01,0.02,0.03,0.04,0.05}
Replacement rate	{0.5,0.6,0.7,0.8,0.9,1}
Vesting Period (months)	{0,1,2,3,4,5,6}
Duration (months)	{2,4,6,8,10,12,∞}

The results in Table 11 suggest that, in general, an unemployment benefit system based on risk-pooling would have modest effects on labor markets, except in the case of very generous replacement rates. The main effects would be a modest increase in the unemployment rate, the reallocation of labor from wage employment into self employment, and an increase in the average wage in both the formal and informal sectors. The unemployment benefit system increases the flow value of unemployment and the flow value of formal jobs (through an increase in wages and a higher expected income when the labor contract ends), but the first effect dominates for a high value of the bargaining power parameter. Because the flow value of unemployment increases, individuals receiving job offers are less likely to take them, which increases the unemployment rate. In all cases simulated, the increase in the unemployment rate as a result of introducing unemployment insurance is below 1 percentage point and in most cases less than or equal to 0.6 percentage points, which is to be compared to an unemployment share of 2.02 percent in the absence of unemployment insurance. As individuals become less likely to take job offers, unemployment rises and the labor market tightness measure $\left(\theta = \frac{V}{U}\right)$ falls. Other things being equal, individuals are then more likely

[28] Changing the value of the bargaining coefficient, of course, changes the employment shares and the distribution of wages. The impacts of the simulated unemployment benefit systems are analyzed against this new (counterfactual) steady state.

Table 7: Effects of Alternative Unemployment Benefit Systems on Employment Shares

Pay Roll Tax	Replace-ment Rate	Vesting Period	Dura-tion	Share Unemployed	Share Informal	Share Formal	Share Self Employed
Estimated Bargaining Power ($\beta = 0.7646$)							
0	0	0	0	0.0%	0.0%	0.0%	0.0%
0.01	0.5	6	**6**	0.4%	-0.5%	-0.3%	0.3%
0.01	0.5	6	**12**	0.4%	-0.7%	0.0%	0.3%
0.01	0.5	**2**	6	0.5%	-0.5%	-0.3%	0.3%
0.01	0.5	**5**	6	0.4%	-0.5%	-0.3%	0.3%
0.01	**0.7**	6	6	0.6%	-2.6%	-0.6%	2.7%
0.01	**0.9**	6	6	1.0%	-1.2%	0.3%	0.0%
0.02	0.7	6	6	0.6%	-2.6%	-0.6%	2.7%
0.05	0.7	6	6	0.6%	-2.6%	-0.8%	2.7%
0	**0.5**	**0**	**0**	1.8%	-0.1%	-1.7%	0.0%
0	**1**	**0**	**0**	45.9%	-23.0%	-26.6%	3.6%
Low Bargaining Power ($\beta = 0.3$)							
0	0	0	0	0.0%	0.0%	0.0%	0.0%
0.01	0.5	6	**6**	-4.3%	-0.5%	12.1%	-7.3%
0.01	0.5	6	**12**	-4.5%	-0.4%	14.1%	-9.2%
0.01	0.5	**2**	6	-4.4%	-0.4%	12.5%	-7.6%
0.01	0.5	**5**	6	-4.3%	-0.5%	12.1%	-7.3%
0.01	**0.7**	6	6	-1.9%	1.6%	6.6%	-6.3%
0.01	**0.9**	6	6	0.3%	1.5%	-1.4%	-0.4%
0.02	0.7	6	6	-2.0%	1.7%	6.6%	-6.3%
0.05	0.7	6	6	-1.1%	2.1%	6.4%	-7.4%
0	**0.5**	**0**	**0**	-6.5%	4.5%	12.8%	-10.8%
0	**1**	**0**	**0**	38.7%	-12.0%	-20.6%	-6.2%

N.B.: Table presents the difference in steady state values between the simulated and baseline scenarii.

Table 8: Effects of Alternative Unemployment Benefit Systems on Average Earnings

Pay Roll Tax	Replacement Rate	Vesting Period	Duration	Mean Wage Informal	Mean Wage Formal	Mean Wage Self Employed
Estimated Bargaining Power ($\beta = 0.7646$)						
0	0	0	0	0.0%	0.0%	0.0%
0.01	0.5	6	**6**	12.3%	4.0%	-4.9%
0.01	0.5	6	**12**	13.4%	3.8%	-4.9%
0.01	0.5	**2**	6	12.3%	4.0%	-4.9%
0.01	0.5	**5**	6	12.3%	4.0%	-4.9%
0.01	**0.7**	6	6	11.3%	4.1%	3.5%
0.01	**0.9**	6	6	22.5%	1.7%	2.0%
0.02	0.7	6	6	11.3%	3.2%	3.5%
0.05	0.7	6	6	11.5%	0.4%	3.5%
0	**0.5**	**0**	**0**	24.2%	8.9%	-7.0%
0	**1**	**0**	**0**	-86.3%	-2.8%	35.8%
Low Bargaining Power ($\beta = 0.3$)						
0	0	0	0	0.0%	0.0%	0.0%
0.01	0.5	6	**6**	0.6%	1.7%	1.2%
0.01	0.5	6	**12**	0.6%	0.9%	1.1%
0.01	0.5	**2**	6	0.6%	1.2%	1.1%
0.01	0.5	**5**	6	0.6%	1.7%	1.2%
0.01	**0.7**	6	6	1.1%	-5.4%	0.7%
0.01	**0.9**	6	6	0.4%	-7.4%	1.2%
0.02	0.7	6	6	1.1%	-5.9%	0.6%
0.05	0.7	6	6	0.4%	-10.3%	0.7%
0	**0.5**	**0**	**0**	0.6%	-18.2%	1.3%
0	**1**	**0**	**0**	-40.0%	109.2%	-1.1%

N.B.: Table presents the difference in steady state values between the simulated and baseline scenarii.

to enter self employment. The share of self employment can increase by up to 3 percentage points in the simulations examined here.

As discussed above, the unemployment benefit system also has a direct impact on wages in the formal sector, increasing the flow value of formal jobs and dampening the negative effect that the increase in the flow value of unemployment has on acceptance rates. In the simulations, the average wage in the formal sector increases by between 3 and 5 percent (more in the case of unemployment assistance). The other important effect is the change in average earnings of self employed workers. Because earnings in that sector only depend on the level of human capital of the individual, changes provide information about the types of workers who are more likely to move in or out of the sector. The results show that average earnings fall when the replacement rate of the unemployment insurance systems is 50%, indicating that low skilled workers disproportionately move to this sector. This also explains the increase in the average wage of the informal sector since those leaving are more likely to be low skilled workers. Moreover, the unemployment benefit system has an additional indirect positive effect on the wages of informal sector workers in that it increases the flow value of unemployment through the increase in formal sector wages. When the replacement rate is above 50%, however, average earnings for the self employed increase, indicating an inflow of workers with higher productivity. The interpretation is that as the replacement rate increases, the additional flow value of unemployment makes self employment attractive even for high skilled workers - for whom the "opportunity cost of unemployment is higher".

Overall, the effects of the unemployment benefit system become more important as the replacement rate increases. However, there are no systematic effects from changes in the duration of benefits or the vesting period. This is likely to be an artifact of the current simulations. Indeed, we have run only one set of random shocks that affect individuals over time. Because individuals can be in different states at time t in different policy simulations, the individual specific shock at time t (which is the same for a given individual across simulations) sometimes determines whether a job offer is made and sometimes whether a job is destroyed. For instance, the share of formal employment can change because initial changes in behavior put individuals on a stochastic path where they receive fewer formal job offers.

The contribution rate also appears to have only a marginal impact on the unemployment rate and employment shares, at least when kept below 5 percent. It does, however, considerably dampen the increase in the average formal sector wage. Indeed, the main effect of the contribution rate is to reduce the wage paid for a given skill level, meaning that firms pass at least part of the cost of additional payroll taxes through to the worker in the form of lower wages. Other things being equal, this reduces the likelihood of taking formal sector jobs but not enough to have substantial effects on employment shares.

Lowering the bargaining power of workers, as seen by Table 12, changes the results considerably. As indicated above, introducing UI reduces the likelihood of individual accepting informal job offers or opportunities for self employment. However, those receiving formal job offers are this time more likely to take them, and formal sector wages rise under UI. When the bargaining power of workers (β) falls, formal sector wages also fall. However, introducing a UI system induces a larger increase in formal sector wages when bargaining power is low than when it is high. This is because the bargaining process puts more weight on the flow value of unemployment (relative to the flow value of employment) when bargaining power is low than when it is high, so a given change in the flow value of unemployment leaders to a larger increase in the formal wage, and thereby the value of formal employment, than in the situation with a higher bargaining power for workers. As a result, we see a larger increase in the share of formal employment and a larger reduction in self employment across the simulations (except in the case of very generous unemployment assistance) when β is lower. This result, although not directly relevant for Malaysia, points to the importance that unemployment benefit systems can have in non-competitive labor markets where employers have more bargaining power relative to workers.

As noted in footnote 7, the model assumes that the state can enforce eligibility rules and ensure that only the unemployed can draw UI benefits. Relaxing this assumption so as to allow those employed in the informal sector (either in wage or self employment) to draw benefits would have both direct and indirect effects on the model's outcomes. The direct effect of allowing UI receipt while in the informal sector is that the value of informal sector jobs and self employment is increased since at least with some probability benefits can be drawn while working. The value of unemployment also increases since part of the benefits can continue to be drawn if the individual exits unemployment into an informal sector job. Another, indirect, effect of relaxing the assumption is that formal sector employment becomes more valuable, as the level of unemployment benefits which enters the value function increases. Presumably, therefore, workers would be less likely to remain unemployed and more likely to work in any sector, shortening unemployment durations. Without formally estimating this more complex model, however, it is impossible to quantify how much the share in each employment state would change as enforcement is weakened.

7 Conclusion

This paper has presented a Diamond-Mortensen-Pissarides type of macro labor market model to analyze the effects of introducing unemployment insurance. Although the application focuses in Malaysia, the results are likely to be relevant for other middle income countries. The model has four labor market states (unemployment, self employment, informal wage employment and formal

wage employment) and was estimated to replicate the structure of the labor market in Malaysia in 2009 and the distribution of earnings for informal, formal and self employed workers. The model is estimated using direct techniques, OLS regression, selection-corrected regression and simulated pseudo method of moments estimators on data from Malaysia in 2009.

The model is then used to simulate the effects of alternative unemployment benefit system designs that depend on the replacement rate, the vesting period for benefits, the duration of benefits, and the contribution rate. The results suggest that introducing an unemployment insurance system in Malaysia would have only a modest negative effect on unemployment if benefits are not overly generous. The model suggests that a UI system would induce a reallocation of labor from wage into self employment while increasing average wages in the formal and informal sectors. The effects on the average earnings of the self employed would depend on the generosity of the system. With a 50% replacement rate, most workers entering self employment would be low skilled workers, driving average earnings down. High skilled workers would move into self employment with more generous systems, thereby increasing average earnings.

Although outcomes appear unaffected by variations in the vesting period needed to obtain benefits, and only slightly affected by changes in the duration of benefits, they do vary with the payroll tax rate, the replacement rate and worker bargaining power. The model suggests that these variations are due to workers changing their behavior in response to changes in the flow value of formal employment (both directly - for those coming from the formal sector - and indirectly for the others) and the flow value of unemployment.

References

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, 07 2009, *119* (539), 1105–1129.

_ , _ , and _ , "Public Sector Employment in an Equilibrium Search and Matching Model," 2010. mimeo.

Bosch, Mariano and Julen Esteban-Pretel, "Cyclical Informality and Unemployment," CIRJE F-Series CIRJE-F-613, CIRJE, Faculty of Economics, University of Tokyo February 2009.

Cunningham, Wendy V., "Unemployment Insurance in Brazil: Unemployment Duration, Wages and Sectoral Choice," Technical Report, World Bank 2000.

Heckman, James J., "Sample Selection Bias as a Specification Error," *Econometrica*, January 1979, *47* (1), 153–161.

Holmlund, Bertil, "Unemployment Insurance in Theory and Practice," *Scandinavian Journal of Economics*, March 1998, *100* (1), 113–41.

Holzmann, Robert and Milan Vodopivec, eds, *Reforming Severance Pay*, Washington, DC: The World Bank, 2011.

_ , Yann Pouget, Milan Vodopivec, and Michael Weber, "Severance Pay Programs around the World: History, Rationale, Status, and Reforms," IZA Discussion Papers 5731, Institute for the Study of Labor (IZA) May 2011.

Margolis, David N., "Unemployment Insurance versus Individual Unemployment Accounts and Transitions to Formal versus Informal Sector Jobs," December 2008. Centre d'Economie de la Sorbonne mimeo.

_ , Lucas Navarro, and David A. Robalino, "A Structural Model of Informality, Self Employment and Job Search," November 2011. World Bank mimeo.

Miller, John H., "Active Nonlinear Tests (ANTs) of Complex Simulation Models," *Management Science*, June 1998, *44* (6), 820–830.

Pissarides, Christopher A., *Equilibrium Unemployment Theory, 2nd Edition*, Vol. 1 of *MIT Press Books*, The MIT Press, June 2000.

Robalino, David A., Eduardo Zylberstajn, and Juan David Robalino, "Incentive Effects of Risk Pooling, Redistributive and Savings Arrangements in Unemployment Benefit Systems: Evidence from a Job-Search Model for Brazil," IZA Discussion Papers 5476, Institute for the Study of Labor (IZA) February 2011.

Tatsiramos, Konstantinos, "Geographic labour mobility and unemployment insurance in Europe," *Journal of Population Economics*, April 2009, *22* (2), 267–283.

van Ours, Jan C. and Milan Vodopivec, "Does reducing unemployment insurance generosity reduce job match quality?," *Journal of Public Economics*, April 2008, *92* (3-4), 684–695.

Vodopivec, Milan, Andreas Wörgötter, and Dhushyanth Raju, "Unemployment Benefit Systems in Central and Eastern Europe: A Review of the 1990s," *Comparative Economic Studies*, December 2005, *47* (4), 615–651.

A Basic Structure of the Model

Our model considers a labor market with salaried and self employed workers in the spirit of Albrecht et al. (2009). Wage workers can produce with firms offering either a formal or an informal contract. That is, firms can create vacancies in the formal or the informal sector. While formal contracts are subject to labor market regulations like payroll taxes τ, informal contracts are not directly affected by labor market policy. It is also assumed that matches in the informal sector produce a flow value of output which is a fraction $\epsilon < 1$ of what the same match would yield with a formal contract.[29] The parameter ϵ can be interpreted for example as the expected cost of a fine for not complying with regulations. Even though we allow for two types of vacancies for wage jobs, we maintain the assumption that search is random. That is, letting ϕ denote the fraction of informal sector vacancies, the effective contact rate for informal and formal vacancies for the worker are $\phi m(\theta)$ and $(1-\phi)\, m(\theta)$, respectively, where $\theta = \frac{\text{Vacancies}}{\text{Unemployment}}$ is a measure of labor market tightness and $m(\cdot)$ is a standard matching function.

Worker participation in the different sectors will be related to, but will not depend exclusively on, the worker's type. We assume an exogenous distribution, $y \sim F(y)$, $\underline{y} \leq y \leq \overline{y}$, of types across workers.

Self employed workers receive an income y equivalent to their type. Opportunities to work in the self employment sector arrive to the unemployed at exogenous Poisson rate α, and employment ends at exogenous Poisson rate λ_0.

We allow for *ex post* idiosyncratic initial match productivity in both wage sectors. When a worker of type y meets a prospective employer with a vacancy, she draws a match-specific productivity, $x \sim G_i(x|y)$, $\underline{x} \leq x \leq \overline{x}$, where the subscript $i \in \{F, I\}$ indicates whether the job is in the formal or informal sector. In order to relate x with y, we assume first-order stochastic dominance, i.e., $y' > y \Rightarrow G_i(x|y') < G_i(x|y)$ like in Albrecht et al. (2010). This means that the higher the value of the worker type indicator y, the greater her expected productivity. Once x is realized, the parties decide to produce if the net surplus of the match is positive and continue searching otherwise. The match surplus, and thereby the negotiated wage, depend both on the productivity x and on the worker's type.

As an attractive feature of the model, note that there is not perfect segmentation of large groups of workers across the sectors given that match productivity not only depends on the worker's type but also on match-specific characteristics. It is thus possible to have workers with employment histories in the three sectors as reflected in the data for many developing countries.

[29] A similar assumption is done in Bosch and Esteban-Pretel (2009).

B Regression Results

Table 9: Results From Estimation of y

VARIABLES	log(Self Employment Income)	Selection Into Self Employment
SPM/SPVM	0.368***	-0.221***
	(0.0289)	(0.0204)
STPM/HSC/STA	0.696***	-0.424***
	(0.0648)	(0.0435)
Sijil	0.578***	-0.383***
	(0.0876)	(0.0647)
Diploma	1.096***	-0.509***
	(0.0592)	(0.0358)
Advanced Degree	1.558***	-0.629***
	(0.0694)	(0.0367)
No Degree	-0.343***	0.239***
	(0.0282)	(0.0209)
Female	-0.646***	-0.140***
	(0.0244)	(0.0252)
Age	0.0762***	0.0158***
	(0.00655)	(0.00473)
Age^2	-0.0950***	0.0144***
	(0.00697)	(0.00556)
Constant	9.249***	-1.913***
	(0.220)	(0.0992)
$atanh\,(\rho)$		-0.783***
		(0.0744)
$\log(\sigma)$		0.0280
		(0.0311)
Observations	63,817	63,817

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

NB: Main and selection models also include controls for 16 states. Selection model includes controls for 10 relations to household head and 5 marital statuses

Table 10: Mincer Regressions for Wage and Salary Sectors

VARIABLES	Formal	Informal
SPM/SPVM	0.263***	0.369***
	-0.01	(0.0143)
STPM/HSC/STA	0.456***	0.668***
	-0.0175	(0.0277)
Sijil	0.493***	0.646***
	-0.0203	(0.0377)
Diploma	0.729***	1.086***
	-0.0124	(0.0203)
Advanced Degree	1.126***	1.351***
	-0.0123	(0.0192)
No Degree	-0.199***	-0.256***
	-0.0129	(0.0156)
Female	-0.268***	-0.311***
	-0.00618	(0.0101)
Age	0.0911***	0.122***
	-0.00203	(0.00271)
Age^2	-0.0911***	-0.129***
	-0.00272	(0.00349)
Constant	7.765***	6.998***
	-0.0379	(0.0526)
Observations	30,563	20,070
R^2	0.486	0.469

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

NB: Models also includes controls for 16 states.

C Simulation Results

Table 11: Effects of Changing System Parameters on Employment Shares and Wages

Pay Roll Tax	Replacement Rate	Vesting Period	Duration	Share Unemployed	Share Informal	Share Formal	Share Self Employed	Mean Wage Informal	Mean Wage Formal	Mean Wage Self Employed
0	0	0	0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.01	0.5	6	**2**	0.2%	-1.2%	0.0%	1.0%	7.0%	3.0%	-3.5%
0.01	0.5	6	**4**	0.4%	-0.3%	-0.4%	0.3%	10.6%	3.7%	-4.9%
0.01	0.5	6	**6**	0.4%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	0.5	6	**8**	0.4%	-1.0%	-0.3%	0.9%	12.8%	4.1%	-5.3%
0.01	0.5	6	**10**	0.5%	-0.5%	-0.3%	0.3%	12.5%	4.0%	-4.9%
0.01	0.5	6	**12**	0.4%	-0.7%	0.0%	0.3%	13.4%	3.8%	-4.9%
0.01	0.5	**1**	6	0.5%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	0.5	**2**	6	0.5%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	0.5	**3**	6	0.5%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	0.5	**4**	6	0.5%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	0.5	**5**	6	0.4%	-0.5%	-0.3%	0.3%	12.3%	4.0%	-4.9%
0.01	**0.6**	6	6	0.6%	-1.4%	-0.1%	0.9%	12.8%	4.6%	-4.0%
0.01	**0.7**	6	6	0.6%	-2.6%	-0.6%	2.7%	11.3%	4.1%	3.5%
0.01	**0.8**	6	6	0.7%	-2.7%	-0.5%	2.5%	17.0%	1.6%	4.9%
0.01	**0.9**	6	6	1.0%	-1.2%	0.3%	0.0%	22.5%	1.7%	2.0%
0.02	0.7	6	6	0.6%	-2.6%	-0.6%	2.7%	11.3%	3.2%	3.5%
0.03	0.7	6	6	0.6%	-2.6%	-0.6%	2.7%	11.3%	2.2%	3.5%
0.04	0.7	6	6	0.6%	-2.6%	-0.6%	2.7%	11.3%	1.3%	3.5%
0.05	0.7	6	6	0.6%	-2.6%	-0.8%	2.7%	11.5%	0.4%	3.5%
0	**0.5**	**0**	**0**	1.8%	-0.1%	-1.7%	0.0%	24.2%	8.9%	-7.0%
0	**0.6**	**0**	**0**	2.6%	-1.1%	-2.5%	1.1%	23.5%	9.1%	0.2%
0	**0.7**	**0**	**0**	3.2%	1.0%	-4.0%	-0.2%	33.3%	6.7%	-3.4%
0	**0.8**	**0**	**0**	3.5%	0.5%	-4.4%	0.3%	38.9%	3.9%	2.2%
0	**0.9**	**0**	**0**	5.0%	-1.2%	-7.3%	3.5%	43.7%	4.3%	10.5%
0	**1**	**0**	**0**	45.9%	-23.0%	-26.6%	3.6%	-86.3%	-2.8%	35.8%

N.B.: Table presents the difference in steady state values between the simulated and baseline scenarii. Simulation using parameters from table 6 and using the estimated bargaining power parameter of $\beta = 0.7646$.

Table 12: Effects of Changing System Parameters on Employment Shares and Wages (Low Bargaining Power)

Pay Roll Tax	Replacement Rate	Vesting Period	Duration	Share Unemployed	Share Informal	Share Formal	Share Self Employed	Mean Wage Informal	Mean Wage Formal	Mean Wage Self Employed
0	0	0	0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.01	0.5	6	**2**	-3.7%	-0.2%	10.4%	-6.5%	0.6%	0.0%	1.3%
0.01	0.5	6	**4**	-4.1%	-0.4%	11.8%	-7.3%	0.7%	1.9%	1.2%
0.01	0.5	6	**6**	-4.3%	-0.5%	12.1%	-7.3%	0.6%	1.7%	1.2%
0.01	0.5	6	**8**	-4.5%	-0.4%	14.1%	-9.2%	0.6%	0.9%	1.1%
0.01	0.5	6	**10**	-4.5%	-0.4%	14.1%	-9.2%	0.6%	0.9%	1.1%
0.01	0.5	6	**12**	-4.5%	-0.4%	14.1%	-9.2%	0.6%	0.9%	1.1%
0.01	0.5	**0**	6	-5.9%	5.2%	11.9%	-11.1%	0.7%	-20.3%	1.1%
0.01	0.5	**1**	6	-4.4%	-0.4%	12.5%	-7.6%	0.6%	1.2%	1.1%
0.01	0.5	**2**	6	-4.4%	-0.4%	12.5%	-7.6%	0.6%	1.2%	1.1%
0.01	0.5	**3**	6	-4.3%	-0.5%	12.1%	-7.3%	0.6%	1.7%	1.2%
0.01	0.5	**4**	6	-4.3%	-0.5%	12.1%	-7.3%	0.6%	1.7%	1.2%
0.01	0.5	**5**	6	-4.3%	-0.5%	12.1%	-7.3%	0.6%	1.7%	1.2%
0.01	**0.6**	6	6	-3.2%	0.8%	9.2%	-6.7%	0.7%	-3.6%	1.6%
0.01	**0.7**	6	6	-1.9%	1.6%	6.6%	-6.3%	1.1%	-5.4%	0.7%
0.01	**0.8**	6	6	-1.0%	2.0%	3.6%	-4.6%	0.6%	-8.2%	0.4%
0.01	**0.9**	6	6	0.3%	1.5%	-1.4%	-0.4%	0.4%	-7.4%	1.2%
0.02	0.7	6	6	-2.0%	1.7%	6.6%	-6.3%	1.1%	-5.9%	0.6%
0.03	0.7	6	6	-1.9%	2.5%	5.5%	-6.2%	0.9%	-9.3%	0.9%
0.04	0.7	6	6	-0.6%	1.5%	4.4%	-5.3%	0.6%	-6.0%	0.5%
0.05	0.7	6	6	-1.1%	2.1%	6.4%	-7.4%	0.4%	-10.3%	0.7%
0	**0.5**	**0**	**0**	-6.5%	4.5%	12.8%	-10.8%	0.6%	-18.2%	1.3%
0	**0.6**	**0**	**0**	-5.9%	7.6%	10.1%	-11.8%	1.0%	-25.7%	1.3%
0	**0.7**	**0**	**0**	-4.8%	6.6%	7.1%	-8.8%	0.4%	-22.4%	1.5%
0	**0.8**	**0**	**0**	-4.2%	4.2%	3.0%	-3.1%	0.1%	-11.1%	1.0%
0	**0.9**	**0**	**0**	-0.2%	-2.5%	-5.0%	7.7%	0.2%	24.6%	2.0%
0	**1**	**0**	**0**	38.7%	-12.0%	-20.6%	-6.2%	-40.0%	109.2%	-1.1%

N.B.: Table presents the difference in steady state values between the simulated and baseline scenarii. Simulation using parameters from table 6 and using the estimated bargaining power parameter of $\beta = 0.3$.